

KLASIFIKASI PENYAKIT DIABETES MELLITUS TIPE 2
MENGUNAKAN *LEARNING VECTOR QUANTIZATION*
(LVQ)

SKRIPSI



UIN SUNAN AMPEL
S U R A B A Y A

Disusun Oleh :

Risti Vika Arvianti

NIM. H72215020

PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS NEGERI SUNAN AMPEL
SURABAYA

2019

PERNYATAAN KEASLIAN

Saya yang bertanda tangan di bawah ini,

Nama : Risti Vika Arvianti

NIM : H72215020

Program Studi : Matematika

Angkatan : 2015

Menyatakan bahwa saya tidak melakukan plagiat dalam penulisan skripsi yang berjudul “KLASIFIKASI PENYAKIT DIABETES MELLITUS TIPE 2 MENGGUNAKAN *LERANING VECTOR QUANTIZATION* (LVQ)”. Apabila suatu saat nanti terbukti saya melakukan tindakan plagiat, maka saya bersedia menerima sanksi yang telah ditetapkan.

Demikian pernyataan keaslian ini saya buat dengan sebenar-benarnya.

Surabaya, 31 Juli 2019

Yang menyatakan,

A handwritten signature in black ink is written over a yellow and green 6000 Rupiah meter stamp. The stamp features the Garuda Pancasila emblem and the text 'METERAI TEMPEL', '6000', and 'RUPIAH'. A serial number 'KCF3AF74437885' is also visible on the stamp.

(Risti Vika Arvianti)

NIM H72215020

LEMBAR PERSETUJUAN DOSEN PEMBIMBING

Skripsi oleh

NAMA : Risti Vika Arvianti

NIM : H72215020

JUDUL : Klasifikasi Penyakit Diabetes Mellitus Tipe 2 Menggunakan *Lerning Vector Quantization* (LVQ)

Ini telah diperiksa dan disetujui untuk diujikan.

Surabaya, 29 JULI 2019

Dosen Pembimbing 1



(Dian C. Rini Novitasari, M.Kom)

NIP.198511242014032001

Dosen Pembimbing 2



(Putroue Keumala Intan, M.Si)

NIP.198805282018012001

PENGESAHAN TIM PENGUJI SKRIPSI

Skripsi oleh

NAMA : Risti Vika Arvianti

NIM : H72215020

JUDUL : Klasifikasi Penyakit Diabetes Mellitus Tipe 2 Menggunakan *Lerning Vector Quantization* (LVQ)

Telah dipertahankan di depan tim penguji skripsi

Pada hari Rabu Tanggal 31 Juli 2019

Penguji I

(Dian C. Rini Novitasari, M.Kom)

NIP.198511242014032001

Penguji II

(Putrout Keumala Intan, M.Si)

NIP.198805282018012001

Penguji III

(Nurissaidah Ulinnuha, M.Kom)

NIP. 19901102201432004

Penguji IV

(Wika Dianita Utami, M.Sc)

NIP. 199206102018012003

Mengetahui,
Dekan Fakultas Sains dan Teknologi
Universitas Jember



(Dr. Leni Purwati, M.Ag)

NIP.196512211990022001



KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL SURABAYA
PERPUSTAKAAN

Jl. Jend. A. Yani 117 Surabaya 60237 Telp. 031-8431972 Fax.031-8413300
E-Mail: perpus@uinsby.ac.id

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika UIN Sunan Ampel Surabaya, yang bertanda tangan di bawah ini, saya:

Nama : RISTI VIKA ARVIANTI
NIM : H72215020
Fakultas/Jurusan : SAINS DAN TEKNOLOGI / MATEMATIKA
E-mail address : ristivikaarvianti@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan UIN Sunan Ampel Surabaya, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah :

☒ Skripsi ☐ Tesis ☐ Desertasi ☐ Lain-lain (.....)

yang berjudul :

KLASIFIKASI PENYAKIT DIABETES MELLITUS TIPE 2 MENGGUNAKAN
LEARNING VECTOR QUANTIZATION (LVQ)

beserta perangkat yang diperlukan (bila ada). Dengan Hak Bebas Royalti Non-Eksklusif ini Perpustakaan UIN Sunan Ampel Surabaya berhak menyimpan, mengalih-media/format-kan, mengelolanya dalam bentuk pangkalan data (database), mendistribusikannya, dan menampilkan/mempublikasikannya di Internet atau media lain secara *fulltext* untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan UIN Sunan Ampel Surabaya, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini yang saya buat dengan sebenarnya.

Surabaya, 2 Agustus 2019

Penulis

(Risti Vika Arvianti)

Namun alangkah baiknya jika melakukan upaya pencegahan suatu penyakit sebelum datangnya penyakit, karena kesehatan yang diberikan Allah kepada hambanya sangatlah mahal harganya dan tidak dapat dibeli dengan uang. Pepatah arab mengatakan bahwasannya mencegah lebih baik dari pada pengobatan.

[illegible]

Menurut *World Health Organization* (WHO) penderita penyakit diabetes di Indonesia akan terus meningkat dari 108 juta jiwa penderita diabetes pada tahun 1980 dan pada tahun 2014 meningkat menjadi 422 juta jiwa. Beberapa provinsi di Indonesia yang terdiagnosis penyakit diabetes yaitu Provinsi Yogyakarta sebanyak (2,6%), Sulawesi Utara sebanyak (2,4%), Provinsi Kalimantan Timur sebanyak (2,3%), Provinsi Sulawesi Tengah sebanyak (3,7%), Provinsi Sulawesi Selatan sebanyak (3,4%), Provinsi Kalimantan Barat sebanyak (11,1%), Provinsi Riau sebanyak (10,4%), Provinsi Nanggroe Aceh Darussalam sebanyak (8,5%),

Adapun faktor-faktor yang mempengaruhi diabetes mellitus tipe 2 menurut penelitian yang pernah dilakukan oleh (Dewi Prasetyani, 2011) terdapat 10 faktor yang dapat mempengaruhi diabetes mellitus tipe 2 yaitu jenis kelamin, status perkawinan, pekerjaan, pola makan, merokok, aktivitas fisik, obesitas, stress, kadar gula darah. Oleh karena itu untuk menghindari penyakit diabetes mellitus tipe 2 terdapat beberapa upaya yang harus dilakukan untuk pencegahan penyakit diabetes yaitu meningkatkan pengetahuan tentang diabetes mellitus, menjaga pola makan dan berolahraga secara cukup, mempertahankan berat badan sesuai dengan umur dan tinggi badan, selalu melakukan cek kadar gula darah (Hasnah, 2009).

[illegible]

Metode LVQ bisa digunakan mengidentifikasi suatu tulisan tangan dengan menggunakan huruf lontara seperti yang dilakukan oleh (Saputra, 2015), LVQ juga bisa digunakan untuk identifikasi pembicara dengan menggunakan ekstraksi ciri *principal component analysis* seperti yang dilakukan oleh (Sugiyanto, 2015), LVQ juga bisa digunakan untuk memprediksi penyakit *Tuberculosis* paru (TB paru) seperti penelitian yang dilakukan oleh (Rahmadani & Jaya, 2018) yang memperoleh hasil akurasi 100%, juga bisa digunakan untuk optimasi vektor bobot *learning vector*

Berdasarkan beberapa penelitian terdahulu yang pernah menggunakan metode klasifikasi *Learning Vector Quantization*. Pada penelitian ini, peneliti tertarik untuk meneliti lebih lanjut untuk mengetahui dan mendalami tentang metode Learning Vector Quantization dengan judul “Klasifikasi Penyakit Diabetes Mellitus Menggunakan *Learning Vector Quantization*”.

Adapun rumusan masalah dari penelitian ini antara lain:

- [illegible]

D. Manfaat Penelitian

Manfaat pada penelitian ini adalah untuk memberikan kontribusi bagi perkembangan pengolahan data medis khususnya pada klasifikasi pada data diabetes mellitus tipe 2 dengan menggunakan metode *Learning Vector Quantization* (LVQ).

D. Manfaat Penelitian

F. Batasan Masalah

BAB II

KAJIAN PUSTAKA

A. Diabetes Mellitus (DM)

Diabetes mellitus (DM) adalah suatu kondisi kadar gula dalam darah meningkat. Diabetes merupakan penyakit gangguan yang berkaitan dengan produksi energi pada sel manusia yang disebabkan karena faktor keturunan, dan makan. Diagnosis khas dari penyakit diabetes mellitus pada umumnya ditandai dengan polyuria (banyak kencing), polydipsia (dehidrasi yang berlebihan), polifagia (selalu merasakan rasa lapar setiap saat), akan tetapi berat badan menurun (Misnadiarly, 2006).

Menurut *International Diabetes Federation*(Federation, 2015)

Diabetes terdapat 3 jenis, yaitu:

1. Diabetes Mellitus tipe 1

Diabetes Mellitus Tipe 1 terjadi karena sistem pertahanan pada tubuh menyerang sel beta yang mana sel beta penghasil insulin pankreas yang dapat mengakibatkan tubuh tidak dapat memproduksi insulin yang dibutuhkan. Penyakit diabetes ini menyerang pada semua kalangan umur. Terutama diabetes mellitus tipe 1 terjadi pada kalangan anak-anak dan remaja. Orang dengan penderita Diabetes Mellitus Tipe 1 sangat memerlukan insulin tiap hari untuk mengendalikan kadar glukosa dalam

Diabetes Mellitus Tipe 2 termasuk kedalam jenis penyakit diabetes yang paling umum diderita oleh orang dewasa, tetapi sangat memungkinkan untuk menyerang pada anak-anak dan remaja.

[illegible]

Gejala penyakit diabetes gestasional biasanya ditandai dengan hiperglikemia saat kehamilan yang langka dan sulit dibedakan dengan gejala kehamilan normal, tetapi mungkin termasuk dehidrasi dan sering buang air kecil. Skrining dengan cara tes toleransi glukosa oral sangatlah dianjurkan. Skrining harus dilakukan diawal kehamilan untuk wanita berisiko tinggi, dan antara minggu 24 dan minggu 28 kehamilan. Wanita dengan hiperglikemia terdeteksi selama kehamilan memiliki risiko besar terhadap tekanan darah yang sangat tinggi dan makrosomia janin (secara signifikan lebih besar dari rata-rata bayi), yang dapat membuat kelahiran vagina sulit dan berisiko.

[illegible]

Terdapat 2 faktor menyebabkan munculnya penyakit Diabetes Mellitus Tipe 2 yaitu faktor yang tidak dapat dimodifikasi dan faktor risiko yang dapat dimodifikasi (Gibney, 2008).

a. Riwayat keluarga (genetik)

[illegible]

c. Konsumsi Karbohidrat Kompleks/Serat

Peningkatan pembuangan glukosa yang dirangsang insulin pada dosis insulin yang telah ditetapkan dapat dilakukan melalui aktivitas fisik (Bazzano LA, Serdula M, 2005).

d. Indeks glikemik dan beban glikemik

Indeks glikemik digunakan untuk meningkatkan kadar glukosa pada darah yang dapat dilakukan dengan membagi tingkatan makanan yang mengandung karbohidrat, roti putih merupakan standar makanan yang digunakan untuk mengontrol kadar glukosa supaya tidak meningkat (Willet W, J Liu s., 2002).

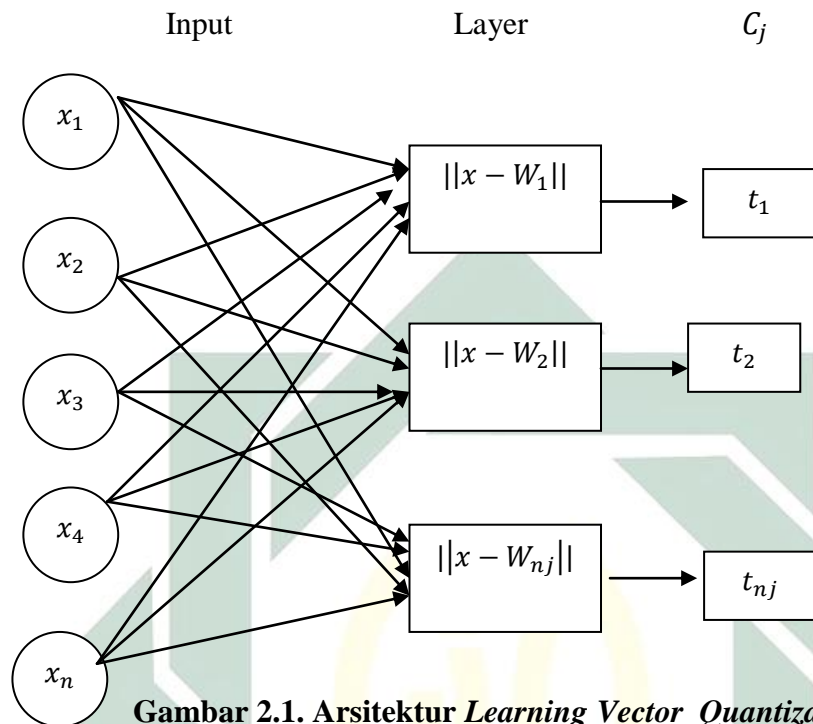
e. Konsumsi Magnesium

Sayuran yang berdaun hijau, biji-bijian dan kacang-kacangan merupakan makanan yang banyak mengandung magnesium digunakan untuk mengurangi resiko diabetes tipe 2. Karena magnesium memiliki hubungan berbanding terbalik dengan kejadian diabetes tipe 2 (Sendih, 2006).

g. Konsumsi Alkohol

D. Learningp Vector Quantization

[illegible]



Gambar 2.1. Arsitektur *Learning Vector Quantization*

Deskripsi variabel dari gambar diatas sebagai berikut :

Tabel 2.1. Definisi variabel

Variabel	Definisi
X	Vektor training ($x_1, x_2, x_3, x_4, \dots, x_n$)
T	Target untuk vektor training sebanyak t_{nj} target yaitu (t_1, t_2, \dots, t_{nj})
W_j	Vektor bobot untuk unit ke-j yaitu ($w_{1j}, w_{2j}, \dots, w_{nj}$)
C_j	Kelas hasil komputasi
$ x - W_j $	Jarak Euclidean antara vektor input dengan vektor bobot untuk unit output ke-j.
W	Vektor bobot
α	<i>Learning rate</i>

Algoritma Learning Vector Quantization (LVQ)

F. Normalisasi

Metode LVQ menggunakan konsep dasar *competitive learning neural networks*, yang mana semua sel akan dilatih untuk membentuk lapisan input, dan semua sel tersebut juga akan mendapatkan output yang sama. Lapisan kompetitif akan secara otomatis belajar untuk mengklasifikasikan vektor-vektor input. Berdasarkan pada konsep “*winner-take-all*”, kelas-kelas yang didapat sebagai hasil dari lapisan-lapisan kompetitif ini hanya bergantung pada jarak antara vektor-vektor input. Vektor input yang mendekati sama maka lapisan kompetitif akan mengklasifikasikan kedua vektor input tersebut ke dalam kelas yang sama (M Indra Halim Arsyadwi Akbari, Astri Novianty S.T., M.T, Casi Setianingsih S.T, 2017).

Normalisasi atau yang biasa disebut dengan proses penskalaan nilai atribut dari data sehingga bisa terletak pada rentang tertentu. Normalisasi data bertujuan untuk menempatkan nilai dalam dalam range 0-1 dengan standarisasi. Untuk menskalakan data dalam jangkauan [0,1] menggunakan persamaan sebagai berikut (Arifin & Sasongko, 2018):

$$\bar{x}_{ik} = \frac{\bar{x}_{ik} - \min(x_k)}{\max(x_k) - \min(x_k)} \quad (2.3)$$

G. Jarak Euclidean

Menurut (Richard A. Johnson, 2016) jarak Euclid merupakan jarak geometris antar dua obyek data. Semakin dekat jarak maka semakin mirip

suatu objek data tersebut. Bentuk umum jarak Euclid (d) dapat diperoleh dengan (Jannah, 2010):

$$\begin{aligned} x &= (x_1, x_2, \dots, x_n) \\ y &= (y_1, y_2, \dots, y_n) \\ d(x, y) &= \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2 + \dots + (y_n - x_n)^2} \end{aligned} \quad (2.4)$$

Jarak *Euclid* memiliki beberapa kelebihan yaitu dapat digunakan untuk mencerminkan ketidaksamaan dua pola, memiliki daya tarik yang intuitif seperti yang umum digunakan untuk mengevaluasi kedekatan objek dalam dua atau tiga dimensi. Menurut (Ghozali, 2016). Jarak *Euclid* juga memiliki beberapa kelemahan diantaranya :

1. Jarak *Euclid* sangat sensitif terhadap besarnya sampel dan besarnya sebaran varians, jika kasus yang sedang dibandingkan memiliki varians yang sangat berbeda maka jarak *Euclid* menjadi tidak akurat. Pengguna jarak *Euclid* menjadi tidak efektif bila antarpeubah terdapat korelasi, oleh karena itu dapat dilakukan analisis komponen utama untuk menghilangkan korelasi antarpeubah.
2. Kecenderungan dari skala terbesar, mendominasi yang lain.

H. Confusion Matrix

Confusion matrix digunakan untuk mengetahui persentase tingkat akurasi dan persentase tingkat error. Berikut confusion matrix ditunjukkan pada Tabel 2.2 berikut:

$$Recall = TP / (TP + FN) \quad (2.8)$$

Presisi dan *recall* dapat diberi nilai dalam bentuk angka dengan menggunakan perhitungan persentase($1 - 100\%$) atau dengan menggunakan bilangan antara 0-1. Sistem rekomendasi akan dianggap baik jika nilai presisi dan *recall*nya tinggi (Powers, 2011).

4. Sensitivitas

$$Sensitivitas = \frac{TP}{(TP+FN)} \quad (2.9)$$

BAB III

METODE PENELITIAN

A. Jenis Penelitian

Penelitian tentang klasifikasi penyakit diabetes mellitus dengan menggunakan *Learning Vector Quantization* (LVQ) termasuk dalam jenis penelitian kuantitatif jika dilihat dari proses pengambilan data dan penyelesaiannya, karena data yang digunakan dalam penelitian ini merupakan data angka. Penelitian kuantitatif merupakan suatu proses penelitian yang menggunakan data berupa angka sebagai alat untuk menganalisis keterangan mengenai apa yang ingin diketahui. Penelitian ini lebih mengedepankan kemampuan dalam mengolah data serta angka-angka dengan menggunakan rumus.

B. Data Penelitian

Data yang digunakan dalam penelitian ini adalah data tentang penyakit diabetes mellitus tipe 2. Data tersebut didapatkan di rumah sakit. Penelitian ini menggunakan metode *Learning Vector Quantization* untuk klasifikasi dengan menggunakan variabel yang terdapat pada Tabel 3.1 dan data masukan yang terdapat pada Tabel 3.2.

1. Proses pembelajaran

2. Proses Pengenalan

[illegible]

BAB IV

PEMBAHASAN

A. Proses Klasifikasi

Pada penelitian ini tahap pertama yang dilakukan untuk melakukan proses klasifikasi dengan analisis data menggunakan variabel usia (x_1) yang berbentuk kontinu, obesitas (x_2) berbentuk numerik, kadar gula darah (x_3) berbentuk numerik, faktor keturunan (x_4) berbentuk nominal, pola makan (x_5) berbentuk nominal. Selanjutnya dari data tersebut di ambil sampel 10 data pertama dari data masukan seperti pada Tabel 4.1.

Tabel 4.1. Normalisasi Data

No	x_1	x_2	x_3	x_4	x_5
1	20	35	45	1	0
2	70	80	320	0	0
3	90	80	400	0	0
4	80	75	500	0	0
5	65	83	330	0	1
6	45	65	60	1	0
7	48	88	227	0	1
8	45	70	88	1	1
9	65	70	250	1	1
10	70	75	30	0	0

Langkah awal dalam menormalisasikan data menggunakan Min-Max yaitu menentukan nilai Maksimal dan nilai Minimal setiap variabel. Pada variabel usia nilai maksimal adalah 88 dan nilai minimal usia adalah 20. Pada variabel obesitas nilai maksimal yaitu 92 dan nilai minimal obesitas adalah 35. Pada variabel kadar gula darah nilai maksimal adalah 510 dan nilai minimal gula darah adalah 30. Selanjutnya dilakukan

perhitungan normalisasi data dengan menggunakan Persamaan (2.3) yang diperoleh hasil pada Tabel 4.2.

Tabel 4.2.Hasil Normalisasi

No	x_1	x_2	x_3	x_4	x_5
1	0	0	0.031	1	0
2	0.735	0.789	0.604	0	0
3	1.029	0.789	0.771	0	0
4	0.882	0.702	0.979	0	0
5	0.662	0.842	0.625	1	1
6	0.368	0.526	0.0625	0	0
7	0.412	0.930	0.410	0	1
8	0.368	0.614	0.121	1	1
9	0.662	0.614	0.458	1	1
10	0.735	0.702	0	0	0

Setelah melakukan proses normalisasi data diubah menjadi bentuk biner [0,1] sesuai kebutuhan pada jaringan. Pengubahan bentuk biner dilakukan dengan ketentuan sebagai berikut:

$$x_i(\text{biner}) = 1 \text{ jika } x \text{ bernilai } > 0,5$$

$$x_i(\text{biner}) = 0 \text{ jika } x \text{ bernilai } < 0,5$$

Sehingga diperoleh hasil dari pengubahan bentuk biner yang dapat dilihat pada Tabel 4.3.

Tabel 4.3. Hasil pengubahan bentuk biner

No	x_1	x_2	x_3	x_4	x_5
1	0	0	0	1	0
2	1	1	1	0	0
3	1	1	1	0	0
4	1	1	1	0	0
5	1	1	1	1	1
6	0	1	0	1	1
7	0	1	0	0	1
8	0	1	0	1	1
9	1	1	0	1	1
10	1	1	0	0	0

Setelah melakukan pengubahan ke dalam bentuk biner selanjutnya dilakukan perhitungan menggunakan metode LVQ secara manual.

B. Contoh Perhitungan dengan LVQ

Pada subbab ini menggunakan perhitungan dengan data yang sebenarnya. Misalnya diketahui 10 input vektor dalam 2 kelas sebagai berikut:

Tabel 4.4.Data Matriks input

No	x_1	x_2	x_3	x_4	x_5	Kelas
1	0	0	0	1	0	Diabetes
2	1	1	1	0	0	Normal
3	1	1	1	0	0	Diabetes
4	1	1	1	0	0	Diabetes
5	1	1	1	1	1	Diabetes
6	0	1	0	0	0	Normal
7	0	1	0	0	1	Diabetes
8	0	1	0	1	1	Diabetes
9	1	1	0	1	1	Diabetes
10	1	1	0	0	0	Diabetes

Tabel 4.5. Data Bobot

No	x_1	x_2	x_3	x_4	x_5	Kelas
5	1	1	1	1	1	Diabetes
6	0	1	0	0	0	Normal

Jarak pada bobot ke 1(1 1 1 1 1

$$= \sqrt{(1-1)^2 + (1-1)^2 + (1-1)^2 + (0-1)^2 + (0-1)^2} = 1.4142136$$

Jarak pada bobot ke 2(1 .005, 0, 0, -0.005, 0)

$$= \sqrt{(1 - 1.005)^2 + (1 - 0)^2 + (1 - 0)^2 + (0 + 0.005)^2 + (0 - 0)^2}$$
$$= 1.4142312$$

Jarak terkecil pada bobot ke-1

Target data ke 2 adalah pada kelas 1 (positif diabetes)

Karena target data ke 2 terdapat pada kelas ke 2, maka w ke-1 baru adalah:

$$W_j(\text{baru}) = W_j(\text{lama}) + \alpha(X_i - W_j(\text{lama}))$$

$$W_{11} = W_{11} + \alpha \times (X_{41} - W_{11}) = 1 + 0.005 \times (1 - 1) = 1$$

$$W_{12} = W_{12} + \alpha \times (X_{42} - W_{12}) = 1 + 0.005 \times (1 - 1) = 1$$

$$W_{13} = W_{13} + \alpha \times (X_{43} - W_{13}) = 0 + 0.005 \times (1 - 1) = 1$$

$$W_{14} = W_{14} + \alpha \times (X_{44} - W_{14}) = 0 + 0.005 \times (0 - 1) = 0.995$$

$$W_{15} = W_{15} + \alpha \times (X_{45} - W_{15}) = 0 + 0.005 \times (0 - 1) = 0.995$$

$$W1 (baru) = (1, \quad 1, \quad 1, \quad 0.995, \quad 0.995)$$

Data ke 3 (1 1 1 0 0)

Jarak pada bobot ke 1 (1, 1, 1, 0.995, 0.995)

$$= \sqrt{(0-1)^2 + (1-1)^2 + (0-1)^2 + (0-0.98507)^2 + (0-0.98507)^2}$$
$$= 1.41443711$$

Jarak pada bobot ke 2 (1.005, 0, 0, -0.005, 0)

$$= \sqrt{(0 - 1.005)^2 + (1 - 0)^2 + (0 - 0)^2 + (0 + 0.005)^2 + (1 - 0)^2}$$
$$= 1.4177623$$

Jarak terkecil pada bobot ke-1

Target data ke 7 adalah pada kelas 1 (positif diabetes)

Karena target data ke 7 terdapat pada kelas 1 maka w ke-1 baru adalah

$$W_j \text{ (baru)} = W_j \text{ (lama)} + \alpha(X_i - W_j \text{ (lama)})$$

$$W_{11} = W_{11} + \alpha \times (X_{71} - W_{11}) = 0 + 0.005 \times (0 - 1) = 1$$

$$W_{12} = W_{12} + \alpha \times (X_{72} - W_{12}) = 1 + 0.005 \times (1 - 1) = 1$$

$$W_{13} = W_{13} + \alpha \times (X_{73} - W_{13}) = 1 + 0.005 \times (0 - 1) = 0.995$$

$$W_{14} = W_{14} + \alpha \times (X_{74} - W_{14}) = 0.98507 + 0.005 \times (0 - 0.98507) \\ = 0.98515$$

$$W_{15} = W_{15} + \alpha \times (X_{75} - W_{15}) = 0.98507 + 0.005 \times (1 - 0.98507) \\ = 0.98515$$

$$W_1 = (1, \quad 1, \quad 0.995, \quad 0.98515, \quad 0.98515)$$

Data ke-8(0 1 0 11)

Jarak terkecil pada bobot akhir W_2 , sehingga data tersebut masuk ke kelas

2 atau normal. Sedangkan pada data asli masuk kedalam kelas diabetes sehingga perlu dilakukan iterasi selanjutnya untuk perhitungan selanjutnya.

5. Data ke 2 (0 1 0 0 0)

kelas 1

$$= \sqrt{(0 - 1)^2 + (1 - 1)^2 + (0 - 0.99003)^2 + (0 + 0.98515)^2 + (0 + 0.98515)^2}$$
$$= 2.93138$$

kelas 2

$$= \sqrt{(0 - 1.00498)^2 + (1 - 0)^2 + (0 - 0)^2 + (0 + 0.01003)^2 + (0 + 0.005)^2}$$
$$= 0.97813 \text{ (output)}$$

Jarak terkecil pada bobot akhir W_2 , sehingga data tersebut masuk ke kelas 2 atau normal. Sedangkan pada data asli masuk kedalam kelas normal sehingga hasil dari lvq dengan data asli sesuai.

C. Hasil Klasifikasi LVQ

Pada penelitian ini proses klasifikasi penyakit diabetes mellitus dengan menggunakan 115 data diabetes mellitus dapat dihitung secara manual dan dapat dihitung menggunakan aplikasi matlab. Untuk mengoptimalkan hasil klasifikasi dalam proses pelatihan, nilai *learning rate* 0.001 sampai 0.009 dan membagi data latih dan data uji menjadi tiga pola yang berbeda yaitu 60% data *training* dan 40% data *testing*, 70% data *training* dan 30% data

testing, dan 80% data *training* dan 20% data *testing*(Hota, Shrivastava, & Singh, 2013)

Tabel 4.7 Hasil Klasifikasi (DTr-60%, DTs-40%) dengan epoch 1000

Training	Testing	LR	TP	FN	FP	TN	Recall (%)	Presisi (%)	Akurasi (%)	Sensitivitas (%0
60%	40%	0.001	44	0	2	0	100	95.65	95.65	100
		0.002	44	0	2	0	100	95.65	95.65	100
		0.003	44	0	2	0	100	95.65	95.65	100
		0.004	44	0	2	0	100	95.65	95.65	100
		0.005	44	0	2	0	100	95.65	95.65	100
		0.006	44	0	2	0	100	95.65	95.65	100
		0.007	44	0	2	0	100	95.65	95.65	100
		0.008	43	1	2	0	97.73	95.56	93.48	97.73
		0.009	44	0	2	0	100	100	95.65	100

Pada Tabel 4.6. di lakukan percobaan dengan menggunakan data training sebanyak 69 data training dan data testing sebanyak 46 data dengan 1000 epoch, pada data testing terdapat 2 pasien normal dan 44 data positif diabetes Pada percobaan dengan menggunakan *learning rate* 0.008 menghasilkan TP =43 menunjukkan jumlah pasien yang positif diabetes dan teridentifikasi positif Diabetes. FN =1 menunjukkan jumlah pasien yang positif Diabetes tetapi teridentifikasi normal. FP =2 menunjukkan jumlah pasien yang normal tetapi teridentifikasi Diabetes, TN =0 menunjukkan jumlah pasien yang normal dan teridentifikasi normal. Recall 97.73% menunjukkan proporsi aktual yang benar diidentifikasi. Presisi 95.56% menunjukkan kecocokan antara permintaan informasi dengan jawaban permintaan tersebut, akurasi 93.48% menunjukkan bahwa klasifikasi penyakit Diabetes mellitus tipe 2 terdapat kemungkinan tepat (benar), dan nilai sensitivitas 97.73% yang menunjukkan bahwa ketika dilakukan uji klasifikasi pada pasien yang positif diabetes dan

pasien yang normal maka terdapat banyak pasien yang positif diabetes. Maka pada model terbaik yang memiliki hasil minimum terdapat *learning rate* 0.008.

Tabel 4.8 Hasil Klasifikasi (DTr-60%, DtS-40%) dengan Epoch 2000

Epoch 2000										
Training	Testing	LR	TP	FN	FP	TN	Recall (%)	Presisi (%)	Akurasi (%)	Sensitivitas (%)
60%	40%	0.001	44	0	2	0	100	95.65	95.65	100
		0.002	44	0	2	0	100	95.65	95.65	100
		0.003	41	3	2	0	93.18	95.35	89.13	93.18
		0.004	44	0	2	0	100	95.65	95.65	100
		0.005	43	1	2	0	97.73	95.56	93.48	97.73
		0.006	43	1	2	0	97.73	95.56	93.48	97.73
		0.007	41	3	2	0	93.18	95.35	89.13	93.18
		0.008	44	0	2	0	100	95.65	95.65	100
		0.009	44	0	2	0	100	95.65	95.65	100

Pada percobaan selanjutnya dilakukan percobaan menggunakan epoch 200 dengan data training 60% dan data testing 40%, pada data testing terdapat banyak data pasien yang positif Diabetes. Pada Tabel 4.8 hasil klasifikasi yang menunjukkan hasil minimum ditunjukkan pada *learning rate* 0.003 dan 0.007, didapatkan sebanyak 41 pasien yang positif diabetes yang teridentifikasi Diabetes (TP), sebanyak 3 pasien yang positif Diabetes yang teridentifikasi normal (FN), terdapat 2 pasien normal yang teridentifikasi Diabetes (FP) dan tidak ada pasien yang normal yang teridentifikasi normal (TN). Sehingga hasil klasifikasi minimum terdapat pada *learning rate* 0.003 dan 0.007

Tabel 4.9 Hasil Klasifikasi (DTr-60%, DtS-40%) dengan Epoch 3000

Epoch 5000										
Training	Testing	LR	TP	FN	FP	TN	Recall (%)	Presisi (%)	Akurasi (%)	Sensitivitas (%0
60%	40%	0.001	44	0	2	0	100	95.65	95.65	100
		0.002	44	0	2	0	100	95.65	95.65	100
		0.003	44	0	2	0	100	95.65	95.65	100
		0.004	43	1	2	0	97.23	95.56	93.48	97.73

Training	Testing	LR	TP	FN	FP	TN	Recall (%)	Presisi (%)	Akurasi (%)	Sensitivitas (%)
		0.005	43	1	2	0	97.23	95.56	93.48	97.73
		0.006	43	1	2	0	97.23	95.56	93.48	97.73
		0.007	44	0	2	0	100	95.65	95.65	100
		0.008	44	0	2	0	100	95.65	95.65	100
		0.009	44	0	2	0	100	95.65	95.65	100

Pada Tabel 4.8 dilakukan percobaan menggunakan epoch 3000 yang terdiri dari 40% data testing dengan *learning rate* 0.001 sampai 0.009. Pada *learning rate* 0.004, 0.005, dan 0.006 didapatkan *recall* 97.23% yang artinya sistem dapat mengklasifikasikan pasien positif Diabetes yang teridentifikasi Diabetes dengan benar, meskipun terdapat sedikit kesalahan dalam proses klasifikasi, yang mana kesalahan tersebut terdapat pada jumlah pasien positif Diabetes yang teridentifikasi normal(FN) yang seharusnya masuk dalam dalam kelas jumlah pasien diabetes yang teridentifikasi Diabetes (TN). Maka model terbaik dengan nilai *recall*, presisi, akurasi, sensitivitas minimum terdapat pada *learning rate* 0.005, 0.006 dan 0.007.

Tabel 4.10. Hasil Klasifikasi (DTr-60%, DtS-40%) dengan Epoch 4000

		Epoch 4000						Recall	Presisi	Akurasi	Sensitivitas
Training	Testing	LR	TP	FN	FP	TN	(%)	(%)	(%)	(%)	
60%	40%	0.001	44	0	2	0	100	95.65	95.65	100	
		0.002	44	0	2	0	100	95.65	95.65	100	
		0.003	44	0	2	0	100	95.65	95.65	100	
		0.004	41	3	2	0	93.18	95.35	89.13	93.18	
		0.005	44	0	2	0	100	95.65	95.65	100	
		0.006	44	0	2	0	100	95.65	95.65	100	
		0.007	44	0	2	0	100	95.65	95.65	100	
		0.008	44	0	2	0	100	95.65	95.65	100	
		0.009	44	0	2	0	100	95.65	95.65	100	

Pada Tabel 4.9 percobaan dilakukan menggunakan epoch 4000 dengan *learning rate* 0.001 sampai dengan 0.0009. Pada percobaan menggunakan *learning rate* 0.004 didapatkan presisi 93.35% yang artinya kecocokan antara

Training	Testing	LR	TP	FN	FP	TN	Recall (%)	Presisi (%)	Akurasi (%)	Sensitivitas (%)
60%	40%	0.001	43	1	2	0	97.73	95.56	93.48	97.73
		0.002	44	0	2	0	100	95.65	95.65	100
		0.003	44	0	2	0	100	95.65	95.65	100
		0.004	43	1	2	0	97.73	95.56	93.48	97.73
		0.005	43	1	2	0	97.73	95.56	93.48	97.73
		0.006	44	0	2	0	100	95.65	95.65	100
		0.007	44	0	2	0	100	95.65	95.65	100
		0.008	44	0	2	0	100	95.65	95.65	100
		0.009	43	1	2	0	97.73	95.56	93.48	97.73

[illegible]

akurasi dan sensitivitas minimum terdapat pada *learning rate* 0.004, 0.005, dan 0.009.

Tabel 4.12. Hasil Klasifikasi (DTr-60%, DtS-40%) dengan Epoch 6000

		Epoch 5000						Recall	Presisi	Akurasi	Sensitivitas
Training	Testing	LR	TP	FN	FP	TN	(%)	(%)	(%)	(%)	
60%	40%	0.001	44	0	2	0	100	95.65	95.65	100	
		0.002	44	0	2	0	100	95.65	95.65	100	
		0.003	44	0	2	0	100	95.65	95.65	100	
		0.004	44	0	2	0	100	95.65	95.65	100	
		0.005	44	0	2	0	100	95.65	95.65	100	
		0.006	44	0	2	0	100	95.65	95.65	100	
		0.007	44	0	2	0	100	95.65	95.65	100	
		0.008	43	1	2	0	97.73	95.56	93.48	97.73	
		0.009	43	1	2	0	97.73	95.56	93.48	97.73	

Pada Tabel 4.11 dilakukan percobaan dengan epoch 6000 menggunakan *learning rate* 0.001 sampai dengan 0.009. Pada *learning rate* 0.008 dan 0.009 didapatkan nilai sensitivitas 97.73% artinya ketika dilakukan proses klasifikasi pasien positif diabetes dengan yang normal maka 97.73% pasien yang terklasifikasi pada positif Diabetes, pada *learning rate* 0.008 dan 0.009 terdapat kesalahan yang mana kesalahan tersebut terdapat pada FN atau jumlah pasien positif Diabetes yang teridentifikasi normal (FN) seharusnya masuk kedalam kelas jumlah pasien Diabetes yang teridentifikasi Diabetes (TP). Sedangkan pada *learning rate* 0.001 sampai dengan 0.007 didapatkan sensitivitas 100% yang artinya pasien positif Diabetes yang teridentifikasi Diabetes terklasifikasi dengan benar. Maka *learning rate* yang disarankan pada pengujian dengan epoch 6000 adalah *learning rate* 0.008 dan 0.007.

Training	Testing	LR	TP	FN	FP	TN	Recall (%)	Presisi (%)	Akurasi (%)	Sensitivitas (%)
60%	40%	0.001	43	1	2	0	97.73	95.56	93.48	97.73
		0.002	43	1	2	0	97.73	95.56	93.48	97.73
		0.003	43	1	2	0	97.73	95.56	93.48	97.73
		0.004	44	0	2	0	100	95.65	95.65	100
		0.005	44	0	2	0	100	95.65	95.65	100
		0.006	44	0	2	0	100	95.65	95.65	100
		0.007	43	1	2	0	97.73	95.56	93.48	97.73
		0.008	44	0	2	0	100	95.65	95.65	100
		0.009	43	1	2	0	97.73	95.56	93.48	97.73

Tabel 4.14. Hasil Klasifikasi (DTr-60%, DtS-40%) dengan Epoch 8000

Training	Testing	LR	TP	FN	FP	TN	Recall (%)	Presisi (%)	Akurasi (%)	Sensitivitas (%)
60%	40%	0.001	44	0	2	0	100	95.65	95.65	100
		0.002	44	0	2	0	100	95.65	95.65	100
		0.003	41	3	2	0	93.18	95.35	89.13	93.18

positif Diabetes tetapi teridentifikasi normal. $FP = 2$ menunjukkan jumlah pasien yang normal tetapi teridentifikasi Diabetes, $TN = 0$ menunjukkan jumlah pasien yang normal dan teridentifikasi normal. Recall 100% menunjukkan proporsi aktual yang benar diidentifikasi. Presisi 94.29% menunjukkan kecocokan antara permintaan informasi dengan jawaban permintaan tersebut, akurasi 94.29% menunjukkan bahwa klasifikasi penyakit Diabetes mellitus tipe 2 terdapat kemungkinan tepat (benar), dan nilai sensitivitas 100% yang menunjukkan bahwa ketika dilakukan uji klasifikasi pada pasien yang positif Diabetes dan pasien yang normal maka terdapat banyak pasien yang positif Diabetes. Sehingga dapat ditarik kesimpulan bahwasannya pada percobaan tersebut apabila dilihat dari hasil akurasi tampak baik, recall 100% yang menunjukkan bahwa sistem dapat mengklasifikasikan data pada kelas positif diabetes dengan benar walaupun terdapat kesalahan yang mana kesalahan tersebut terletak pada FP yang artinya terdapat 2 pasien normal yang teridentifikasi Diabetes (FP) yang seharusnya masuk pada jumlah normal teridentifikasi normal (TP), dan presisi menunjukkan ketepatan antara data asli dengan hasil klasifikasi. Sehingga *learning rate* yang model terbaik dengan nilai minimum terdapat pada pengujian dengan menggunakan *learning rate* 0.001 sampai 0.009.

Tabel 4.18 Hasil Klasifikasi (DTr-80%, DTs-20%) dengan 1000-10.000 Epoch

Training	Testing	Epoch	LR	TP	FN	FP	TN	Recall (%)	Presisi (%)	Akurasi (%)	Sensitivitas (%)
80	20	1000	0.001	21	0	2	0	100	91.30	91.30	100
			0.002	21	0	2	0	100	91.30	91.30	100
			0.003	21	0	2	0	100	91.30	91.30	100
			0.004	21	0	2	0	100	91.30	91.30	100
			0.005	21	0	2	0	100	91.30	91.30	100

Training	Testing	Epoch	LR	TP	FN	FP	TN	Recall (%)	Presisi (%)	Akurasi (%)	Sensitivitas (%)
			0.006	21	0	2	0	100	91.30	91.30	100
			0.007	21	0	2	0	100	91.30	91.30	100
			0.008	21	0	2	0	100	91.30	91.30	100
			0.009	21	0	2	0	100	91.30	91.30	100

Pada Tabel 4.17. dilakukan percobaan menggunakan data training sebanyak 92 data dan data testing sebanyak 23 data dengan 1000-10.000 epoch yang mana pada data testing terdapat 2 pasien normal dan 21 data positif diabetes. Pada percobaan menggunakan *learning rate* 0.001 sampai dengan 0.009 menghasilkan TP =21 menunjukkan jumlah pasien yang positif Diabetes dan teridentifikasi diabetes. FN =0 menunjukkan jumlah pasien yang positif Diabetes tetapi teridentifikasi normal. FP =2 menunjukkan jumlah pasien yang normal tetapi teridentifikasi Diabetes, TN =0 menunjukkan jumlah pasien yang normal dan teridentifikasi normal. *Recall* 100% menunjukkan proporsi aktual yang benar diidentifikasi. Presisi 91.30% menunjukkan kecocokan antara permintaan informasi dengan jawaban permintaan tersebut, akurasi 91.30% menunjukkan bahwa klasifikasi penyakit Diabetes mellitus tipe 2 terdapat kemungkinan tepat (benar), dan nilai sensitivitas 100% yang menunjukkan bahwa ketika dilakukan uji klasifikasi pada pasien yang positif Diabetes dan pasien yang normal maka terdapat banyak pasien yang positif Diabetes. Sehingga dapat ditarik kesimpulan bahwasannya pada percobaan tersebut apabila dilihat dari hasil akurasi tampak baik, hasil *recall* baik yang menunjukkan bahwa sistem dapat mengklasifikasikan data pada kelas positif diabetes dengan benar walaupun terdapat kesalahan yang mana kesalahan tersebut terletak pada FP yang artinya terdapat 2 pasien normal

BAB V
PENUTUP

A. Simpulan

Berdasarkan hasil dan pembahasan mengenai penelitian klasifikasi penyakit diabetes mellitus tipe 2 menggunakan learning vector quantization dapat disimpulkan bahwa:

1. Proses normalisasi data menggunakan metode Min-Max langkah awal yang harus dilakukan adalah menentukan nilai Minimal dan Maksimal setiap variabel. Pada penelitian ini didapatkan nilai minimal pada variabel usia sebesar 20 dan nilai maksimal usia yaitu 88, nilai minimal pada variabel obesitas yaitu 35 dan nilai maksimal obesitas yaitu 92, nilai minimal pada variabel kadar gula darah adalah 30 dan nilai maksimal kadar gula darah adalah 510. Selanjutnya dilakukan proses normalisasi dan hasilnya dilakukan pembulatan sehingga data akhir berbentuk biner.
2. Proses klasifikasi pada penyakit diabetes mellitus tipe 2 dilakukan dengan menginputkan 5 data dan diperoleh 10 layer. Selanjutnya menghitung bobot tiap layer tersebut dengan rumus jarak *Euclidean*, nilai jarak *Euclidean* terkecil akan dipilih menjadi output.
3. Hasil klasifikasi menggunakan metode *learning vector quantization* dengan data testing sebanyak 40% dilakukan

DAFTAR PUSTAKA

- [illegible]

- Handayani, S. A. (2003). *Faktor-faktor Resiko Diabetes Mellitus Tipe 2 Di Semarang dan sekitarnya*. Universitas Diponegoro Semarang.
- Hasnah. (2009). Pencegahan Penyakit Diabetes Mellitus Tipe 2. *Media Gizi Pangan*, VII(1), 1–4.
- Hota, H. S., Shrivasa, A. K., & Singhai, S. K. (2013). Artificial Neural Network Decision Tree and Statistical Techniques Applied for Designing and Developing E-mail Classifier. *International Journal of Recent Technology and Engineering*, (16), 2277–3878.
- Isnaini, N., & Ratnasari, R. (2018). Faktor risiko mempengaruhi kejadian Diabetes mellitus tipe dua. *Jurnal Kebidanan Dan Keperawatan Aisyiyah*, 14(1), 59–68. <https://doi.org/10.31101/jkk.550>
- Jannah, U. (2010). *Perbandingan Jarak Euclid dengan Jarak Mahalanobis pada Analisis Cluster Hirarki*. 76. Retrieved from <http://etheses.uin-malang.ac.id/id/eprint/6739>
- Juddin, D. R. (2017). *Hubungan Tingkat Pengetahuan Faktor Risiko DM dengan Status DM pada Pegawai Negeri Sipil UIN Alauddin Makassar Tahun 2017*. *Skripsi. Fakultas Kedokteran dan Ilmu Kesehatan*. 10(2014), 63–71.
- Kholis, I. (2016). *Analisis Variasi Parameter Learning Vector Quantization Artificial Neural Network*. 1–12.
- M Indra Halim Arsyah Dwi Akbari, Astri Novianty S.T., M.T, Cati Setianingsih S.T, M. . (2017). Analisis Sentimen Menggunakan Metode Learning Vector Quantization Sentiment Analysis Using Learning Vector Quantization Method. *E-Proceeding of Engineering*, 4(2), 2283–2292.
- Misnadiarly. (2006). *Diabetes Mellitus, Ulcer, Infeksi, Mengenali gejala, Menanggulangi, dan Mencegah Komplikasi*. Jakarta: Pustaka Obor Populer.
- Mutoharoh. (2017). *Pengaruh Pendidikan Kesehatan terhadap Tingkat Pengetahuan tentang Penyakit Diabetes Mellitus pada Penderita Diabetes*

Mellitus Tipe 2 di Desa Ngadiwarno Sukorejo Kendal. 1–65.

Nagi, D. (2006). The Role of Physical Activity in the Prevention of Type 2 Diabetes. *Exercise and Sport in Diabetes: Second Edition*, 54(January), 67–76. <https://doi.org/10.1002/0470022086.ch4>

Nurkhozin, a., Irawan, M. I., & Mukhlash, I. (2011). Klasifikasi Penyakit Diabetes Mellitus Menggunakan Jaringan Syaraf Tiruan Backpropagation Dan Learning. *Prosiding Seminar Nasional Penelitian, Pendidikan Dan Penerapan MIPA*, (7), 1–8.

Ojo, O. (2014). Diabetes in Ethnic Minorities in UK: The Role of Diet in Glucose Dysregulation and Prevalence of Diabetes. *Journal of Food & Nutritional Disorders*, 02(02). <https://doi.org/10.4172/2324-9323.1000110>

PERKENI. (2011). *Konsensus Pengelolaan Diabetes Mellitus Tipe 2 di Indonesia*.
Semarang: PB PERKENI.

POWERS, D. M. . (2011). Estimation of high affinity estradiol binding sites in human breast cancer. EVALUATION: FROM PRECISION, RECALL AND F-MEASURE TO ROC, INFORMEDNESS, MARKEDNESS & CORRELATION. *Journal of Machine Learning Technologies*, 2(1), 37–63.

Puspitaningrum, D. (2006). *Pengantar Jaringan Saraf Tiruan* (F. S. Suyantoro, Ed.). Yogyakarta: C.V ANDI OFFSET.

Putri, N. R. (2012). *Universitas Indonesia Learning Vector Quantization Dengan Logika Fuzzy Untuk Pengenalan Wajah Fuzzy*.

Rahmadani, A. W., & Jaya, A. I. (2018). Prediksi penyakit tuberculosis paru (tb paru) menggunakan metode learning vektor quantization (lvq). *Jurnal Ilmiah Matematika Dan Terapan*, 15, 20–27.

Ramadhan, M. (2017a). *Faktor yang Berhubungan dengan Kejadian Diabetes Mellitus di RSUP DR WAHIDIN SUDIROHUSODO DAN RS Universitas Hasanuddin Makassar Tahun 2017*. Makassar.

- Ramadhan, M. (2017b). *Skripsi Ini Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana Kesehatan Masyarakat*. 1–113.
- Richard A. Johnson, D. W. W. (2016). *Applied Multifariate Statistical Analysis*.
- Saputra, S. S. (2015). *Metode Learning Vector Quantization Pada Jaringan Syaraf Tiruan untuk Mengidentifikasi Tulisan Tangan Huruf Lontara*.
- Soegondo, S. (2007). *Diagnosis dan klasifikasi Diabetes Mellitus terkini. Dalam Penatalaksanaan Diabetes Mellitus terpadu*. Jakarta: Balai Penerbit FKUI.
- Sri Kusumadewi, S. H. (2006). *NEURO-FUZZY: Integrasi Sistem Fuzzy & Jaringan Syaraf*. Yogyakarta: Graha Ilmu.
- Sugiyanto, E. (2015). *Penerapan learning vector quantization untuk identifikasi pembicara dengan menggunakan ekstraksi ciri principal component analysis endrik sugiyanto*. Bogor.
- Widyawati, K., Setiawan, B. D., & Adikara, P. P. (2017). Optimasi Vektor Bobot Learning Vector Quantization Menggunakan Algoritme Genetika untuk Penentuan Kualitas Susu Sapi. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (J-PTIIK) Universitas Brawijaya*, 2(1), 217–225. Retrieved from <http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/767>